AD NUMBER

AD249324

LIMITATION CHANGES

TO:

Approved for public release; distribution is unlimited.

FROM:

Distribution authorized to DoD only; Administrative/Operational Use; OCT 1960. Other requests shall be referred to Air Force Cambridge Research Labs., Bedford, MA.

AUTHORITY

AFCRL ltr 3 Nov 1971

AFCRL -TN-60-1141

APPLICATIONS OF TUNNELING TO ACTIVE DIODES

N. Holonyak, Jr.

General Electric Company
Advanced Semiconductor Laboratory
Syracuse, New York

Scientific Report No. 2b

AF 19(604) - 6623

October, 1960

Prepared for

ELECTRONICS RESEARCH DIRECTORATE
AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
AIR FORCE RESEARCH DIVISION (ARDC)
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS

AFCRL -TN-60-1141

APPLICATIONS OF TUNNELING TO ACTIVE DIODES

N. Holonyak, Jr.

General Electric Company
Advanced Semiconductor Laboratory
Syracuse, New York

Scientific Report No. 2b

AF 19(604) - 6623

October, 1960

Prepared for

ELECTRONICS RESEARCH DIRECTORATE
AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
AIR FORCE RESEARCH DIVISION (ARDC)
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS

Requests for additional copies by Agencies of the Department of Defense, their contractors, and other Government agencies should be directed to the:

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA

Department of Defense contractors must be established for ASTIA services or have their "need-to-know" certified by the cognizant military agency of their project or contract.

All other persons and organizations should apply to the:

U. S. DEPARTMENT OF COMMERCE
OFFICE OF TECHNICAL SERVICES
WASHINGTON 25, D.C.

TABLE OF CONTENTS

I. Abstract

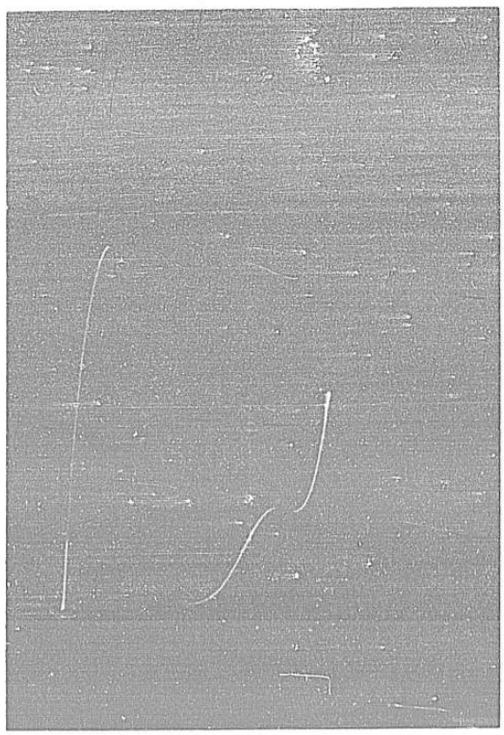
This report covers the results obtained to date in studies of tunneling due to states (bands) in the "forbidden gap" of GaAs and InP, describes a reliability - failure property in GaAs tunnel diodes which bears some resemblance and possibly relationship to formation of "forbidden gap" states, outlines a method of regrowing and doping GaAs (and potentially other compounds) from an alloy-vapor state, and presents some initial results in growing various intermetallic compounds (including epitaxial growth of GaAs) via reaction with a halogen in a closed quartz tube.

A theoretical and experimental study of tunnel diode characteristics and their relationship to circuit performance has been undertaken. The circuits that have been studied are as follows: general purpose logic, sequential circuits, pulse generators and drivers, memory, and active R-C filters. Results to date indicate that the practical applications of tunnel diodes are limited.

II. States (Bands) in the Forbidden Gap of Degenerate GaAs and InP-Secondary Tunnel Currents and Negative Resistances

In early work on tunnel diodes in germanium and silicon a number of workers observed excessive "valley" currents or "humps" in various portions of the I-V characteristics. Longo[1] has reported data on germanium tunnel diodes which exhibit a "hump" and a second negative resistance near the thermal region of the I-V characteristic. Holonyak, et al[2] have reported a similar hump in silicon tunnel diodes. In this report there are described even more striking manifestations of these effects in degenerate GaAs and InP which indicate that states (bands) exist in the forbidden gaps of these materials.

Figure 1 shows the I-V characteristic of a GaAs tunnel diode (at 78°K) which exhibits a secondary tunneling region, with a peak current of 30+ ma at 1.1+ volts, and a second negative resistance region (with faint oscillations visible) near the thermal portion of the I-V characteristic. Secondary tunneling regions as shown in Fig. 1 have been obtained reproducibly in GaAs doped very heavily ($> 10^{19}$ impurity atoms/cm^3) with cadmium and a comparable or slightly greater amount of zinc, and upon which a tunnel junction is formed by alloying to the degenerate p-type substrate pure tin or sulphur-doped tin. The peak in current which occurs at low voltage is due to the well known Esaki component of current[3], whereas the second peak is



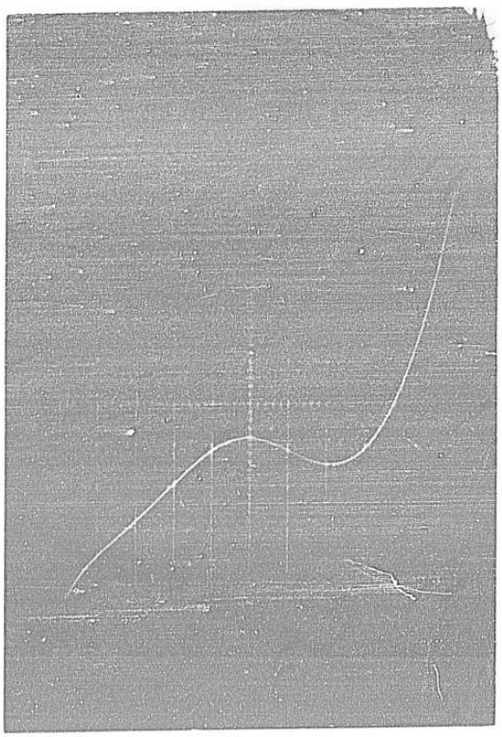
Scale: Horiz. 0.2 VOLTS/DIV.

Vert. 20 MA/DIV.

FIG. 1. Gallium arsenide tunnel diode I-V characteristic at 78°K. Parent crystal doped with cadmium - zinc, alloy-regrown region doped with tin.

attributed to tunneling of electrons from the n-type side of the junction to a band which lies 1 to 1.1 eV above the valence band "edge" on the p-type side of the junction and which may be from 0.1 to 0.3 eV wide, perhaps merging with a "smeared" conduction band edge. It is not considered likely that this band is formed on the n-type (alloy) side of the junction because only properly doped p-type wafers display the effect, i.e. sufficiently heavily doped wafers but not so heavily doped as to cause complete "smearing" of the band structure. Also, in the case of wafers doped by diffusion, it has been possible to reduce the relative magnitude of the secondary tunnel region (peak at 1.1 volts) by etching the junction more and by thereby removing more of the heavily diffused p-type layers, i.e. the heavier doped layers near the surface which are most responsible for the secondary tunneling.

In addition to the secondary tunneling at 1.1 volts which may be observed in GaAs, a strong secondary tunnel region is observed in some units in the voltage range from 0.4 to 0.5. This secondary tunnel current, which occurs in what corresponds to the valley region of a normal GaAs tunnel diode, has been observed in initially fabricated units (i.e. freshly made units which have not undergone life-test) and has been caused to occur in a number of "good" units by life-testing with bias conditions sufficient to operate the diodes in the thermal region. Figure 2 shows the I-V characteristic of a latter



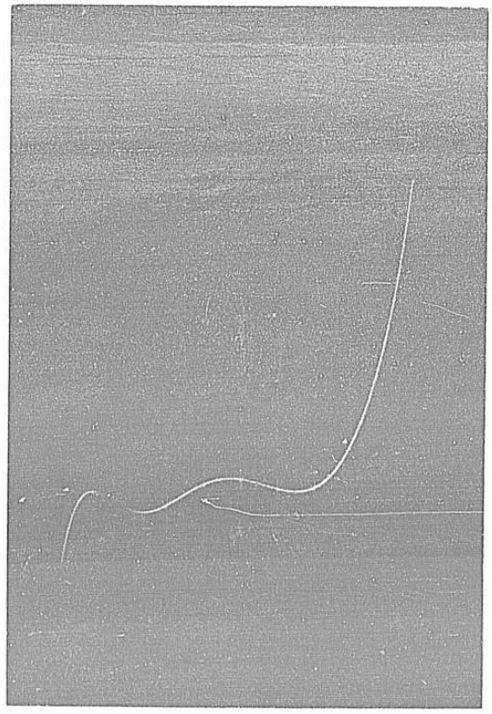
Scale: Horiz. 0.1 VOLTS/DIV.

Vert. 1.0 MA/DIV.

FIG. 2. Gallium arsenide tunnel diode I-V characteristic after life-testing with bias in the thermal region. Esaki component of current almost eliminated; secondary tunnel current peak at 0.5 - volts.

such unit after it has undergone life-testing and has all but completely lost the Esaki component of current, and has developed instead a large secondary tunnel current, 4+ ma at 0.5 - volts. Prior to life-testing (continuous operation with bias in the thermal region) the unit whose characteristic is shown in Fig. 2 displayed a peak current of 10 ma and a valley current less than 1 ma ($I_p/I_V > 10:1$). A similar characteristic in a freshly-made unit is shown in Fig. 3. The unit whose characteristic is shown in Fig. 3 was fabricated upon a p-type wafer which is believed to have possessed appreciable oxygen in the crystal.

In addition to life-tested units that displayed the effects described above, many units have been tested which display other undesired changes in the I-V characteristics. Generally in all life-tested units there is observed an increase in valley current, there is often observed a decrease in peak current (Esaki component of current), there is generally observed a shift of the thermal region towards lower voltages, and quite frequently there is generated the secondary tunnel current ("hump") observed at 0.4 to 0.5 volts. In units which upon life-testing exhibit a decreased Esaki current and a secondary tunnel current at 0.4 to 0.5 volts, an ion drift mechanism could well account



Scale: Horiz. 0.1 VOLTS/DIV.

Vert. 0.5 MA/DIV.

FIG. 3. Gallium arsenide tunnel diode I-V characteristic showing Esaki current and secondary tunnel current. Crystal believed to be contaminated with oxygen.

for the change in the I-V characteristic. The fact that all GaAs units exhibit changes in their characteristics when life-tested indicates the existence of a strong mechanism which could be ion drift, contamination with materials such as oxygen or copper, mechanical failure in the junction region due to local hot-spots or free Ga, or perhaps some unknown mechanism peculiar to GaAs.

Along with the secondary tunneling currents observed in GaAs, similar results have been obtained in a series of InP tunnel diodes fabricated upon degenerate p-type wafers (polycrystalline) doped by diffusion with zinc or a combination of cadmium and zinc. Tunnel junctions were fabricated by alloying onto degenerate p-type InP wafers an indium-tin alloy doped either with tellurium, selenium, or sulphur. Figure 4 shows the I-V characteristic and conductance dI/dV of an InP tunnel diode ($78^{\circ}K$) and is representative of the type of results which were obtained. In the range from -0.2 to $\sim$0.2 volts the I-V characteristic appears like that of a backward diode, i.e. the low-voltage forward (Esaki) component of tunnel current is absent. However, beyond 0.2 volts appreciable tunneling is noted, and a peak in tunnel current occurs at 0.48 volts followed by 0.2+ volts of negative resistance region. These results are quite surprising,

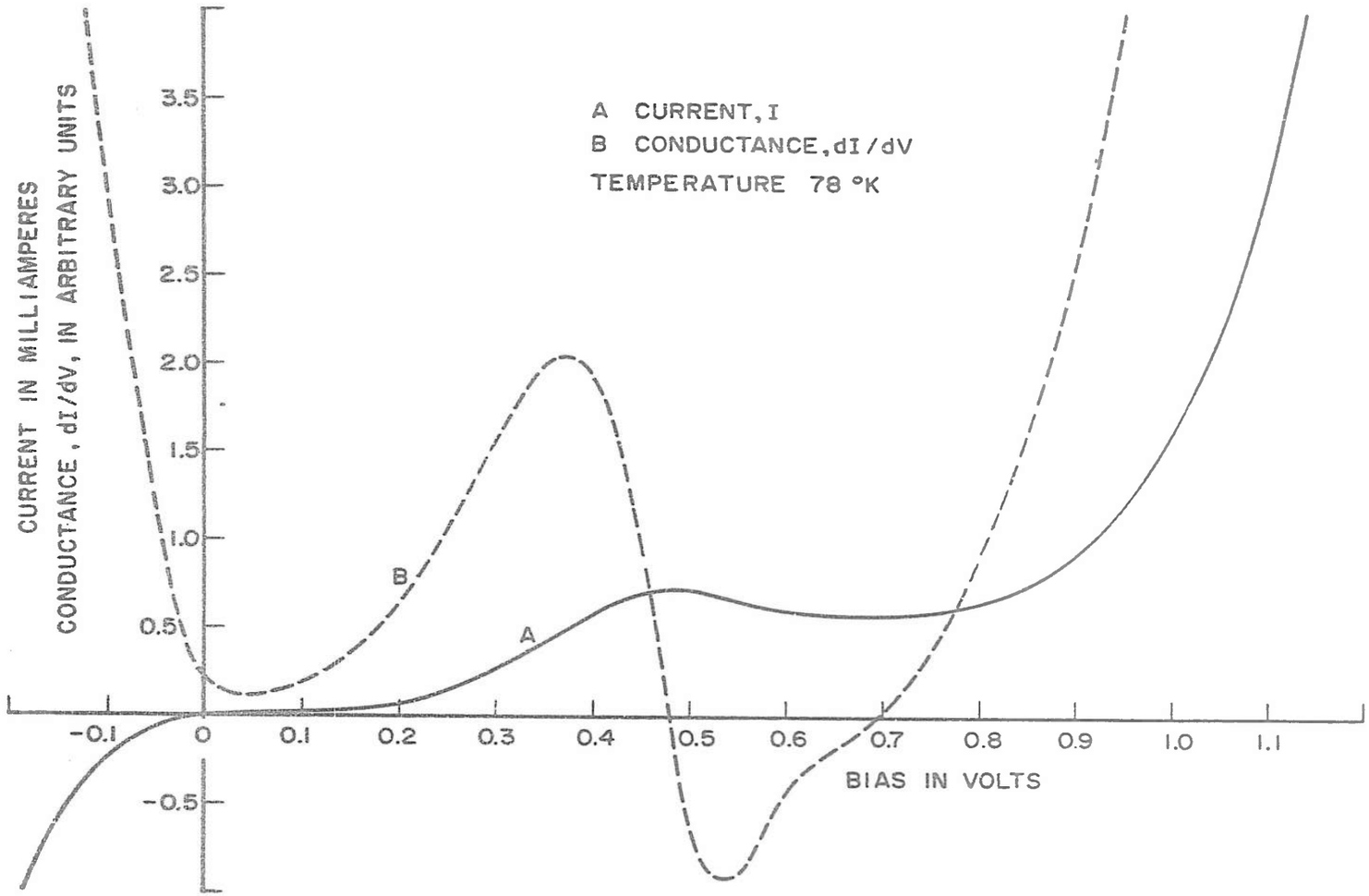


FIG. 4. Indium phosphide tunnel diode I-V characteristic (at 78°K) which exhibits no Esaki component of current.

as it would be expected that ordinarily a tunnel current should occur at much lower voltages and a valley region should exist where in fact in the present case there exists a strong tunnel current.

These data indicate that the p-type diffused wafers were just barely degenerate and that the electrons, available in the degeneracy region of the n-type (alloy) side of the junction, tunnel at higher voltage to a band on the p-type side which is centered 0.4 to 0.5 volts above the "smeared" valence band edge. The exact width and shape of this band is not known since the degree of degeneracy on the n-type side and its effect on the tunnel current are not known. It is almost certain that a band, as described on the p-type side of the junction (in the "forbidden gap"), does not exist on the n-type alloy-side of the junction, because this would require that tellurium, sulphur, and selenium all dope to the same level - a level in this case just sufficient to produce degeneracy. Also, it has been possible to vary the impurity diffusion conditions sufficiently to cause heavier p-type doping; in this case a lower voltage component (Esaki component) of current also can be identified. This component of tunnel current and one as shown in Fig. 4 add to give an overall room temperature I-V characteristic which has the usual monotonically decreasing magnitude of conductance (dI/dV) typical of most tunnel diodes. At liquid nitrogen temperatures the

Esaki component and the secondary component of tunnel current are sufficiently separated to cause the conductance to exhibit an added minimum in the range (variable) from 0.1 to 0.2 volts.

Because of the early nature of this work, it is not certain what factors are most responsible for formation of bands in the "forbidden gap" of degenerate semiconductors. There is good reason to believe that at high doping levels the acceptor and donor elements themselves, at least in part, are responsible for "gap states," perhaps by introducing interstitial, cluster, or vacancy states.

The results presented above concerning secondary tunnel currents lead to several conclusions. First of all, it is probably incorrect to regard a degenerate semiconductor in the general case as possessing a true forbidden gap. Rather, the "forbidden gap" should be considered as a region of reduced density of states with the possibility of localized or smeared regions of higher densities of states that can lead to secondary tunnel currents and also to large excess currents. This work leads to the practical implication that eventually it may be possible to control the density of states in the "forbidden gap" of degenerate semiconductors and as a result make possible tunnel diodes with linear or undulating negative resistance regions which extend over greater voltages, perhaps to the point where thermal (injected) currents become important. This would, in effect, allow greater utilization of the "forbidden gap" of a degenerate semiconductor .

III. Vapor-Alloy Regrown GaAs

A very convenient method of doping semiconductors for fabrication into tunnel diodes is by diffusion[4]. However, doping via diffusion has the disadvantage of giving an impurity concentration which is adequately high at the surface but which is lesser at greater and greater depths below the surface. This introduces undesired series resistance in a tunnel diode and also gives a variation in characteristics (current density) which is a function of the depth to which the junction is alloyed into the wafer. Hence, it is generally preferred to dope uniformly tunnel diode material in the growing process itself.

A very convenient means of accomplishing this in GaAs, and possibly in other intermetallic compounds, is by growing the doped material (degenerate) from an alloy of the impurity and the material itself. For example, an arbitrary quantity of GaAs may be sealed in a 1 cm diameter quartz tube of approximately 6 inch length with 100 mg or more cadmium and may be heated for several hours (or conveniently overnight) to a temperature between 900°C and 1200°C. The large quantity of cadmium in the closed system attacks and forms an alloy with GaAs. If the closed quartz tube is then cooled locally, cadmium begins to condense on the cooled region and thus causes the cadmium vapor pressure over the alloy to drop. Cadmium is then rejected from

the melt, and GaAs finally nucleates and grows with a quantity of cadmium in solid solution corresponding to the solid solubility at the alloy temperature, a temperature which can be held fixed and independent of the conditions needed for freezing of the melt (alloy).

One immediate conclusion of the foregoing is that the above method of doping could be quite useful to establish the solid solubility of cadmium and other impurities in GaAs at various temperatures. Unfortunately the work performed to date was not sufficiently refined to allow growth of large doped crystals and solid solubility measurements were not undertaken. Large masses consisting of large crystallites were grown and were doped either with pure cadmium or cadmium-zinc. Tunnel junctions were alloyed upon selected crystallites, and consistently there were observed I-V characteristics with peak to valley current ratios in the range from 40:1 to 70:1, voltage swings from 1.1 to 1.2 volts, capacitances per milliampere as low as 0.1$\mu\mu$f., and voltages (V_p) at the peak current from 100 to 125 millivolts. In some cases voltages (V_p) at the peak current were observed as low as 60 to 70 millivolts. These are some of the better GaAs tunnel diodes which have been reported to date; unfortunately they along with all others change characteristics when life-tested, and this remains as perhaps the most serious problem extant in GaAs.

IV. Growth of Intermetallic Compounds Via (Halogen) Vapor Transport

Recently Marinace[5] has described a system in which iodine reacts in a closed quartz tube with a germanium crystal and transports and deposits germanium epitaxially upon a lower temperature seed germanium crystal. If the source crystal is doped properly or if a doping impurity is included in the closed tube, it is possible to grow an expitaxially deposited layer of conductivity type opposite to that of the seed crystal and thus form a junction layer, or indeed a layer of almost any conductivity type and resistivity. The question immediately arises whether the same type of results could be expected in intermetallic compounds. The answer is not immediately obvious, as it is not certain whether epitaxial growth would occur or whether the compound would be decomposed in reacting with a halogen.

Experimentally it has been established that a number of intermetallic compounds and mixed compounds may be grown epitaxially. The experiments demonstrating this have been performed in 1 cm diameter closed quartz tubes ranging in length from 6 to 12 inches. A source consisting of one or two intermetallic compounds is sealed into one end of a quartz tube with 5 to 10 mg. of iodine or a similar amount of $MgCl_2$, $SnCl_2$, $ZnCl_2$, $HgCl_2$, $CdCl_2$, $AlCl_2$, or $SbCl_3$, with 1 to 10 mg. of the desired doping impurity, and in some cases with a seed crystal in the opposite end of the closed tube. After sealing under high vacuum the tube is

inserted in a furnace with the source end of the tube heated to a higher temperature than the seed end of the tube. In order to prevent simple thermal decomposition of the crystals, either a small amount of arsenic or phosphorus, depending upon the crystal, is also sealed into each tube. In the heating process iodine or chlorine (from one of the salts) attacks the source crystal, transports the material to the cooler end of the tube, deposits the material on the seed crystal or on the cooler regions of the closed tube, and then repeats the cycle. No attempt has been made to establish conclusively that the crystalline material is transported as the compound or is transported element by element and then re-reacted as the compound. Nevertheless, under proper heating conditions crystalline material is transported and seeded as the compound.

Large crystals, polycrystalline, of GaAs have been grown by this method by seeding directly onto one end of the quartz reaction tube. Iodine has been used for transporting GaAs we well as $ZnCl_2$, $HgCl_2$, $CdCl_2$, $SnCl_2$, $AlCl_2$, and $MgCl_2$. When one of the chloride salts is employed, it is obvious that along with the necessary halogen a doping impurity is present. $ZnCl_2$ and $CdCl_2$ have been used to regrow GaAs and dope it degenerately p-type. Tunnel diodes have been alloyed on the materials so prepared and have yielded peak to valley current ratios as high as 30:1. Crystals vapor regrown with $HgCl_2$ or $MgCl_2$ were p-type but not degenerate.

Crystals vapor regrown with $SnCl_2$ were n-type and p-type when free zinc was employed with $SnCl_2$. When iodine was used as a carrier, n-type and p-type crystals were grown by doping with Te and Zn, respectively.

Figure 5 shows a cross-section of a polycrystalline GaAs p-n junction which was grown by growing, as described above, an n-type GaAs crystal into one end of a quartz tube and by then re-sealing in the quartz tube zinc and growing on the n-type material p-type material. Figure 5 shows a longitudnal cross-section of the crystal (etched). The etch attacked selectively and clearly reveals the two sides of the junction. Electrodes have been affixed to portions of sections such as shown in Fig. 5 and exhibit typical junction I-V characteristics.

Figure 6 shows a cross-section of an n-type GaAs seed upon which has been grown epitaxially p-type GaAs. The junction around the seed is clearly visible. X-ray measurements indicate the deposited material has the orientation of the seed crystal. The usual junction properties are exhibited by such structures. Thus far only slight signs of tunneling have been evident in GaAs epitaxially-grown junctions. Work is in progress to produce epitaxially-grown tunnel junctions to provide a means of junction fabrication other than alloying. This is

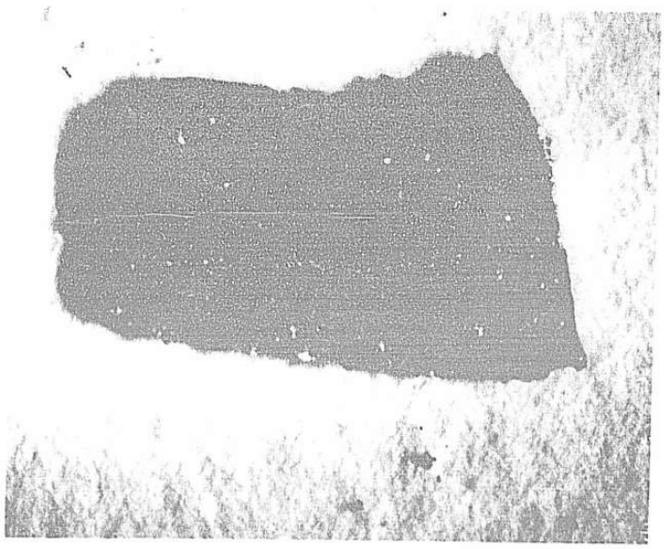


FIG. 5. Gallium arsenide polycrystalline vapor-grown p-n junction. Polycrystalline n-type region grown by means of halogen transport of GaAs; p-type region grown on the n-type material by same process and with zinc enclosed in the system.

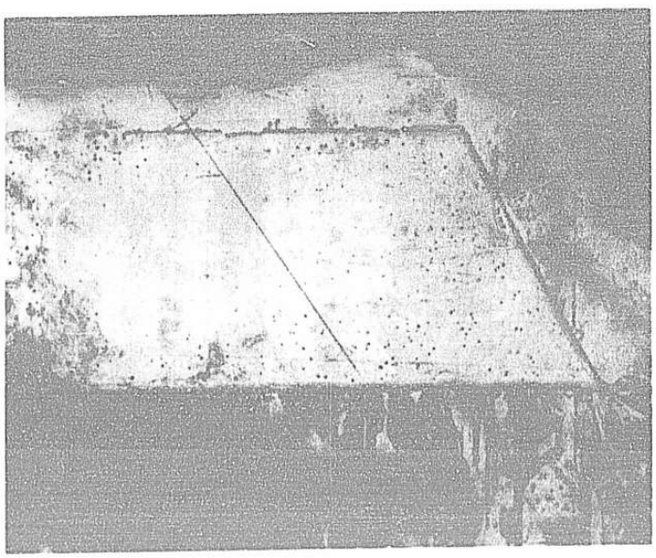


FIG.6. Gallium arsenide p-n junction. N-type seed with p-type GaAs grown epitaxially on seed.

expected to be of value in its own right and also for studies related to failure and change in the characteristics of GaAs tunnel diodes.

Besides the work with GaAs which has been described above, a number of other intermetallic compounds have been regrown by vapor transport. These include GaP, GaSb, and InP. GaP has been transported with iodine and with chlorine as provided by $CdCl_2$ or $SnCl_2$ sealed in the reaction tube. Doping has been accomplished by including in the system free Te, Cd, or Zn. It should be pointed out that small pieces of GaP, essentially powder, may be used as a source crystal and may be seeded and grown ontc the cooler end of a reaction tube as massive polycrystals. If more refined seeding and growing processes are employed, it whould be possible to grow large single crystals of GaP.

In addition to GaAs and GaP, GaSb and InP have been grown by halogen vapor transport. Iodine or $SbCl_3$ have been used to grow GaSb, and $SnCl_2$ has been used to grow InP.

Perhaps one of the most important uses of crystal growing via vapor transport with a halogen is to prepare mixed intermetallic compounds. That is, the source crystals in this case may be two distinct intermetallic compounds such as GaAs and GaP which are to be combined as Ga(AsP) by means of simultaneous vapor transport and seeding as the mixed crystal. This has been successfully

accomplished with GaAs and GaP to yield Ga(AsP) consisting of various proportions of As and P. Hysell has measured the bandgap of one such crystal as being 1.85 eV and estimates the ratio of As to P in the crystal as 6:4. This compares with a weighed ratio of As and P of 5.9:4.1 in the source GaAs and GaP. The crystals produced so far are not homogeneous but are sufficiently uniform to allow optical measurements. Polished specimens (1.85 eV bandgap) of 0.010" thickness visually may be seen to transmit red light. Junctions have been alloyed upon p-type Ga(AsP), prepared as described, and in several cases have yielded tunnel diodes with negative resistances.

V. Measurements

The purpose of the measurement program is to obtain information that can be used to characterize the tunnel diode. This information includes the d-c characteristics V_p, I_p, V_v, I_v, their variations with temperature, the series resistance R_s, and the junction capacitance C. Several different measuring techniques were used during this part of the program.

The first quantity to be measured is R_s, the series resistance of the tunnel diode. This quantity is of interest because it determines the maximum frequency of operation. R_s is of importance in C.W. applications but is not too significant for switching circuits.

The technique for measuring R_s is to drive the tunnel diode to many times the peak current value in the reverse direction and to determine the slope of the I-V characteristic at this point. The reverse current level should be so high that the I-V curve is a straight line. To reach this region of operation requires that the current be 10 to 100 times the peak current of the diode. Currents of this magnitude can easily damage or destroy the diode. This measurement must be made on a pulse basis. As a result no meaningful measurements of R_s have been made.

The capacitance of the diode is a very important parameter in both switching circuits and in C-W circuits. Consequently primary importance was attached to making this measurement.

Because of the problem of stabilizing the tunnel diode in a test circuit a switching load line was used to bias the diode for the purpose of making the capacitance measurements. This technique allows the measurement of the sum of junction and distributed capacitances over a small range of voltages around the valley point. The test circuit is shown in Fig. 7.

The R.F. choke L_1 is used to decouple the bias network from the bridge. Resistor R_3 supplies a current somewhat greater than the diode peak current and R_2 is used to adjust the intercept point on the device characteristics so that the bridge measures a parallel positive resistance of 1 K ohm or greater. Inductor L_2 is used to offset the bridge if capacitances greater that 20 pf are to be measured. The measuring signal across the tunnel diode must be kept to less than 10 millivolts to prevent errors due to the non-linearity of the diode chracteristics. A measuring frequency of 20 Mc was used.[6]

With the exception of three units with capacitances in the range of 50 to 100 pf all the diodes measured had total capactances (C junction + C strays) of 20 pf or less. Four ten milliampere diodes

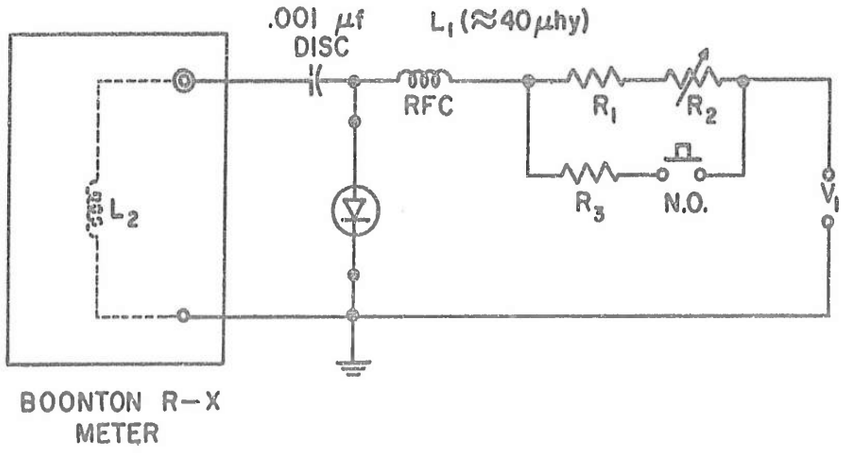


FIG. 7 Capacitance Test Set

had capacitances of 4.6, 3.4, 5.4, and 4.6 pf while the two 20 Ma diodes which were measured had capacitances of 3.8 and 3.6 pf respectively.

The d-c characteristics of a diode V_p, I_p, V_v, and I_v and their variations with temperature are required for any circuit design. These parameters were all measured simultaneously by tracing the characteristic curve of the diode at all the temperatures of interest.

Although this technique appears to be relatively straightforward some difficulties have arisen. Care must be exercised to prevent the diodes from oscillating in the test jigs. For high current diodes with low capacitance oscillations will still occur in the negative resistance region. These oscillations may not adversely affect the measurement of the desired parameters. Fig. 8 shows the I-V characteristic of a low capacitance tunnel diode with 20 ma I_p.

The primary problem has been to display or record the characteristic in such a manner as to allow convenient measurement. An x-y recorder was used because it provided a permanent record. It was found that small errors in the position of the paper in the recorder could mask small variations in the characteristic.

To make these measurements a pair of reference lines must be provided. These lines are the zero voltage and zero current lines.

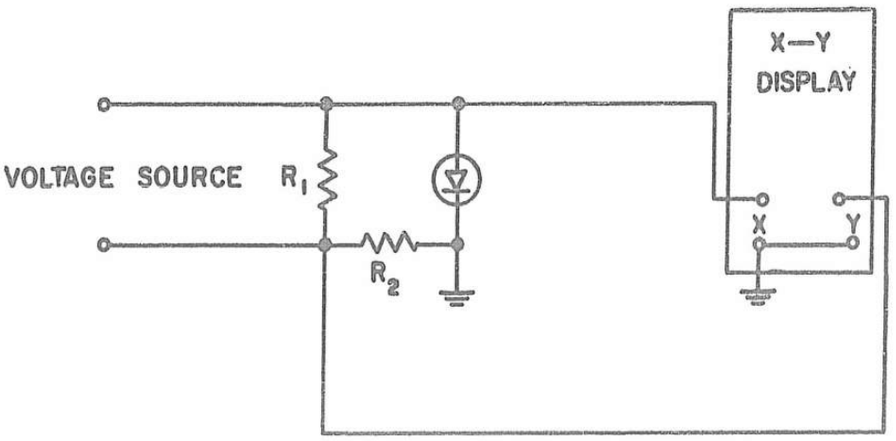
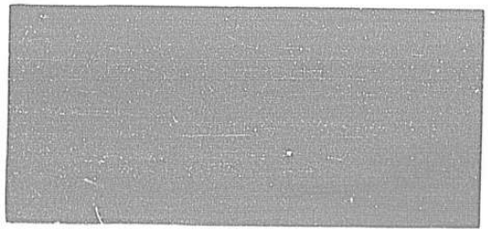


FIG. 8 Curve Tracer

The use of expanded scales is also necessary at the critical points. More sophisticated techniques are available for using very small a-c signals to detect the peak and valley points. These techniques require more complex equipment and so simpler methods are desirable. An evaluation of both methods is currently in progress.

The measurements of the d-c characteristics are made at discrete temperatures over the range of interest. The proposed temperatures are -50°C, -25°C, 0°C, 25°C, 50°C, 75°C, and 100°C. The test setup is illustrated in Fig. 9.



I-V Characteristic of 20ma Tunnel Diode

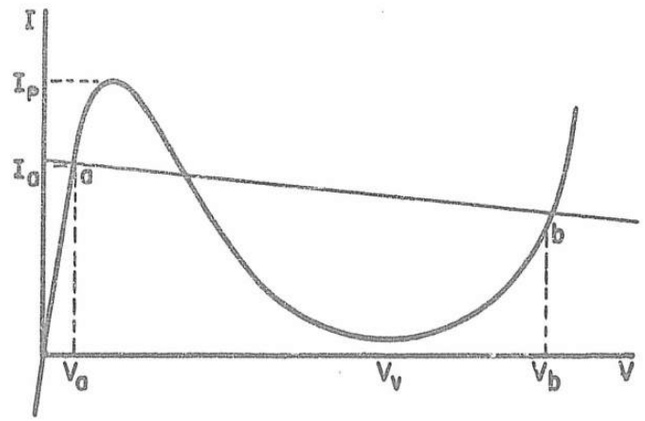
Figure 9

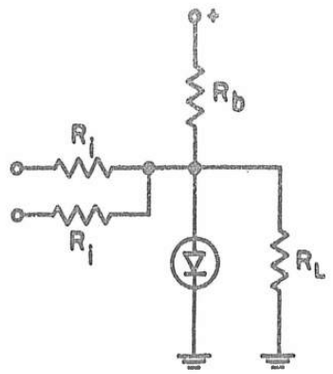
VI. General Purpose Logic

The effect of component tolerances on the performance of tunnel diode logic circuits has been analyzed.[7] The analyses predict that with reasonable diode and component tolerances the number of inputs and outputs is restricted to a value too low for practical use. Recent work has indicated that even this conclusion was optimistic because of some of the assumptions and also the omission of temperature effects. An attempt was made to find new approaches to the problem.

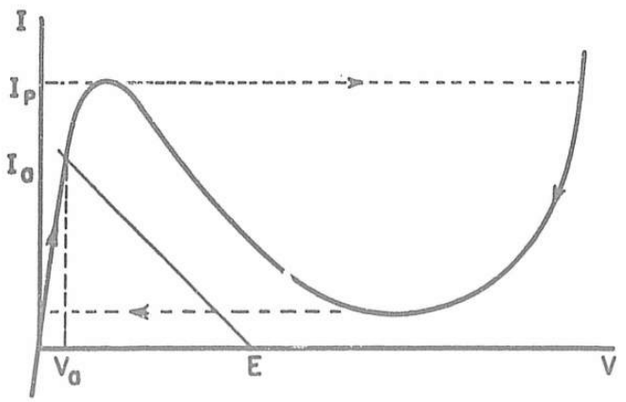
There are two ways of connecting the logic stages depending upon the form of the binary information. The first means is to represent a one by a plus unit voltage and a zero by zero voltage (or very small plus voltage). The second method is to represent a one by a plus unit voltage and a zero by a minus unit voltage, or vice-versa. In the first case logic is performed by the analog addition of the input signals. The function to be performed is determined by the setting of the input threshold. In the second case an odd number of inputs must be used. Logic is also performed by the analog summation of the inputs; however, this sum will be either net positive or negative, the polarity determining the action. This is majority logic.

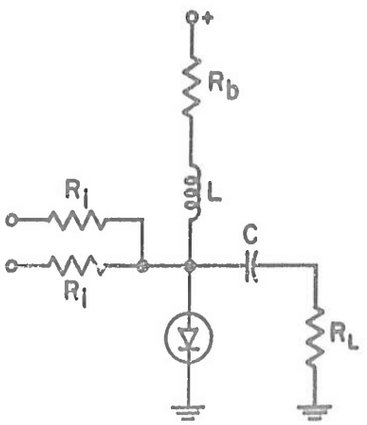
Two examples of analog threshold gates are shown in Fig. 10. These are bistable and monostable gates respectively. If the threshold is exceeded the tunnel diode switches to the high voltage state





BISTABLE





MONOSTABLE

FIG. 10 Analog Threshold Circuits

presenting a "one" output. This output is in the form of a d-c level for the bistable circuit and a pulse for the monostable circuit. In the case of the bistable circuits a reset must be supplied to return all the diodes to the zero state at the end of a logic operation. For high speed logic this pulse must be narrow and of sufficiently high power to reset the diodes. Satisfactory reset operation has not been obtained for logic operating at rates in excess of 5 mc. There is another problem associated with the reset. All logic stages preceding a storage element must reach steady state before they are reset. As a result the bit rate is determined by the propagation delay and the number of levels of logic between storage elements. Timing is also a factor with the monostable circuit since the signals are in the form of pulses. Time coincidence must exist between all inputs to a logic stage.

The most serious problem with analog threshold circuits is the lack of isolation. The input and output circuits of each stage are tied together at a common point. This is a low impedance point until the diode switches at which time the voltage rises. Current can then flow back into a preceding stage causing possible logic malfunctions.

It is possible to achieve logical isolation by means of rectifying diodes; however, the forward voltage drop across the rectifying diode is of the same order of magnitude as the available logic swing. It was anticipated that the higher voltage swing of the gallium arsenide

diode would allow the use of rectifying diodes but the voltage swing is only twice that of germanium and is still comparable to the drop across the rectifying diode. As a result the voltage available to drive the next input would be reduced substantially causing any minor variations that existed to become a much larger percentage of the available signal. This makes the tolerance problem very severe.

Backward diodes, which have a much lower voltage drop than do ordinary rectifying diodes appear to offer the best solution to this problem. If backward diodes are used gallium arsenide tunnel diodes are preferable to germanium because of the higher voltage swing. The variations in the forward voltage drop of backward diodes must be assessed before any conclusions can be reached on their effectiveness.

Logical isolation can also be achieved for these circuits by the use of an inversion stage as shown in Fig. 11. The inputs to the inversion stage are from the elevated outputs of the "and" gates. Resistor R_2 is chosen such that before either "and" gate is switched there is no voltage difference across R_o. When D_1 is switched into the high voltage region current from the upper "and" gate flows only into D_3. When both D_1 and D_2 are switched the combined output switches D_3 and the output of the inversion stage is reduced to "o" with no possibility of backward flow into the "and" gates. This forms a "NAND" circuits.

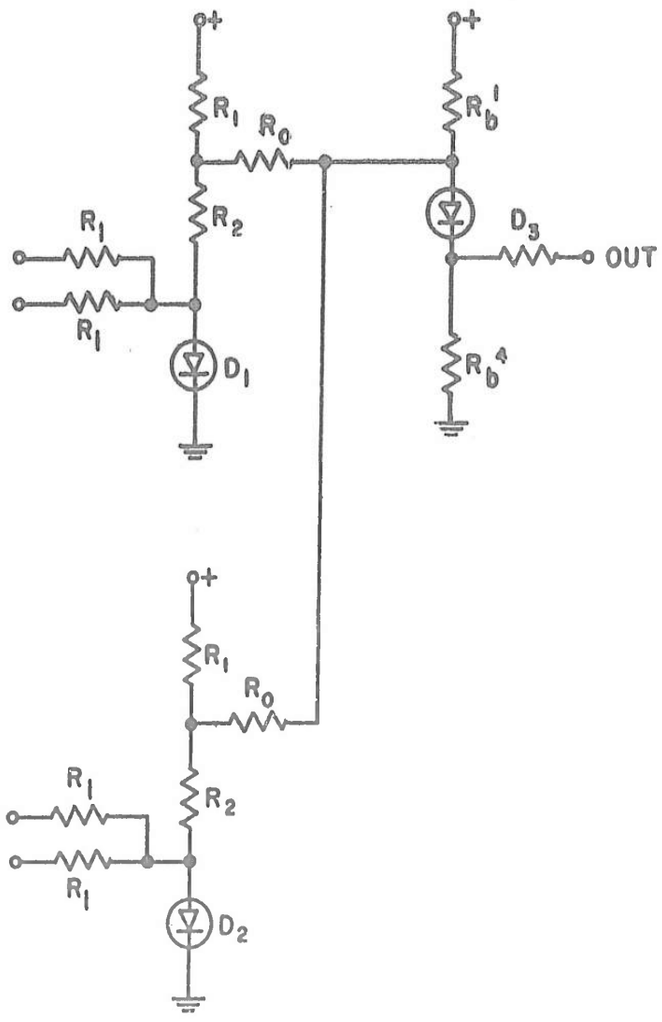


FIG. 11 Logical Isolation With Inversion Stage

This circuit requires that the resistance values be carefully chosen. Even so it is more susceptible to component variations than the conventional circuitry since it depends upon voltage drops across resistors. Another disadvantage of this circuit is that the low power gain of the inverter stage allows it to drive only one other stage. This can be remedied by putting an additional logic stage on the output; however, this can reintroduce the isolation problem. This technique does not appear promising.

"Goto pair" or majority logic operated with a three-phase clock does not need any logical isolation. Since the logic element is always switched one way or the other. The backward flow of signal cannot change the state of a diode already switched into the high voltage region.

It can be shown that "Goto pair" logic must have either a three phase clock to perform logic properly or some form of temporary storage element between logic stages. The use of the three phase clock leads to a more predictable realization, especially at higher speeds because of the discharge characteristics of the temporary storage elements. In the three phase clock scheme if the logic circuits are symmetrical with respect to the inputs and outputs a change in the phase rotation of the clock source causes a change in the direction of information flow.

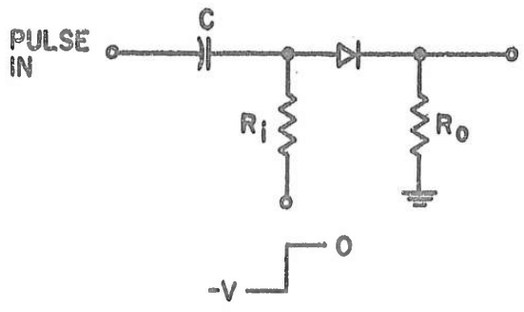
The major problem with "Goto pair" logic has been to develop a satisfactory three phase clock for operation at higher speeds, that is 1 mc or above. The clock is in the form of three square waves (or clipped sine waves) operating at a frequency of twice the bit rate. The three waves are shifted 120^{o} with respect to each other. The generation of such a clock involves high frequency power oscillators as the clock is the logic power supply. This has not been done successfully at speeds in excess of 1 mc.

Another technique for constructing such a clock has been to use a d-c supply with a set of three trigger and reset pulses. This approach has not been successful to date but with the development of high-power fast drivers it shows some promise.

Another approach that has been taken is to use "Goto pair" logic at low speeds where the clock is not a problem. The primary justification for this is that the size and power levels are less than for comparable transistor logic.

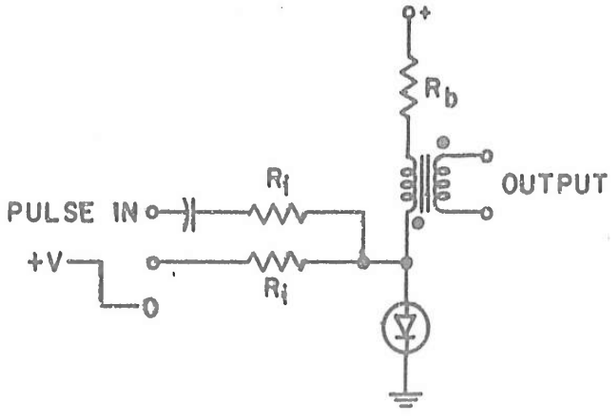
The present effort has been devoted to the study of high speed circuitry since in general transistor and diode logic circuitry or even magnetic circuitry works very well at the lower speeds. At the present time the tunnel diode is not attractive for general purpose logic at these speeds because of the isolation problem or the clock problem. A new approach is required to make tunnel diode logic practical.

Although general purpose tunnel diode logic is impractical some special purpose circuitry has been developed which utilizes the special characteristics of tunnel diodes. One very useful circuit is the high frequency transmission gate or gated pulse generator. Transmission gates as shown in Fig. 12 are commonly used in logic systems to gate pulses. At frequencies in excess of 1 mc the R-C time constant must be so low as to adversely effect the performance of the gate. By using a tunnel diode monostable circuit with two inputs as shown in Fig. 12 a high speed gate can be built. Gates of this type have been operated at speeds up to 25 mc. The only disadvantage is the small output amplitude available.



Conventional Transmission Gate

(a)



(b)

Tunnel Diode Pulse Gate

FIG. 12

VII. Sequential Circuits

Sequential circuits consist of counters, shift registers, and memory elements. It is possible to build either synchronous (clocked) systems or asynchronous systems.

The synchronous systems depend upon a clock supply such as the three phase clock, although it is possible to use any number of phases. Since flip-flops can be built from "and" gates, "or" gates, and inverters it is possible to use conventional "Goto pair" logic elements to construct any type of flip-flop. By using the appropriate form of flip-flop, counters or shift registers can be built; however, this method requires far more logic elements than do other techniques.

Shift registers can be built in a simple manner using a clock. These registers require one or two diodes per bit per phase. No transient storage is required if a three phase clock is used; however, it is required if a single phase clock is used. It is also possible to use a two phase clock scheme without transient storage if rectifying diodes are used for isolation.

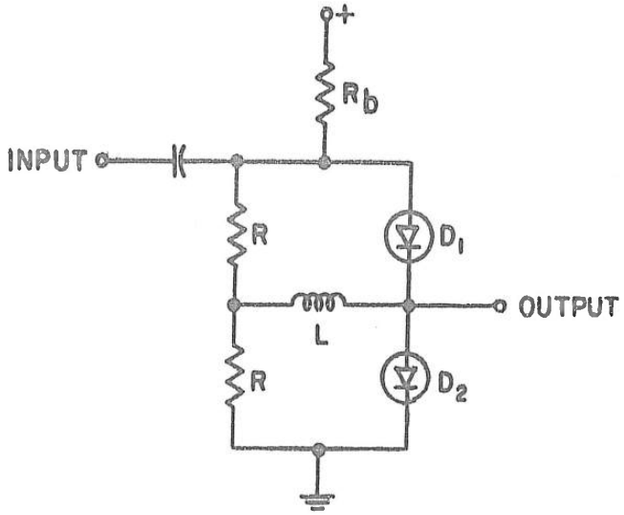
High speed operation of these circuits is restricted because of the problem of generating the clock signals. In the one phase system there is also the problem of charging and discharging the transient storage elements. Because of their simplicity and low power consumption these circuits are attractive for use even at low speeds.

The asynchronous flip-flops are usually constructed of two diodes as shown in Fig. 13. These circuits are primarily counters of scalers although the circuit of Fig. 13b is a tunnel diode analog of the transistor flip-flop except for the lack of complementary outputs. The use of such circuits in a shift register requires the provision of transient storage between stages. It is customary to use the shift pulse to reset all stages to zero. If a stage was storing a "one" the shift or reset pulse would then switch the next stage to the "one" state at the end of the reset pulse.

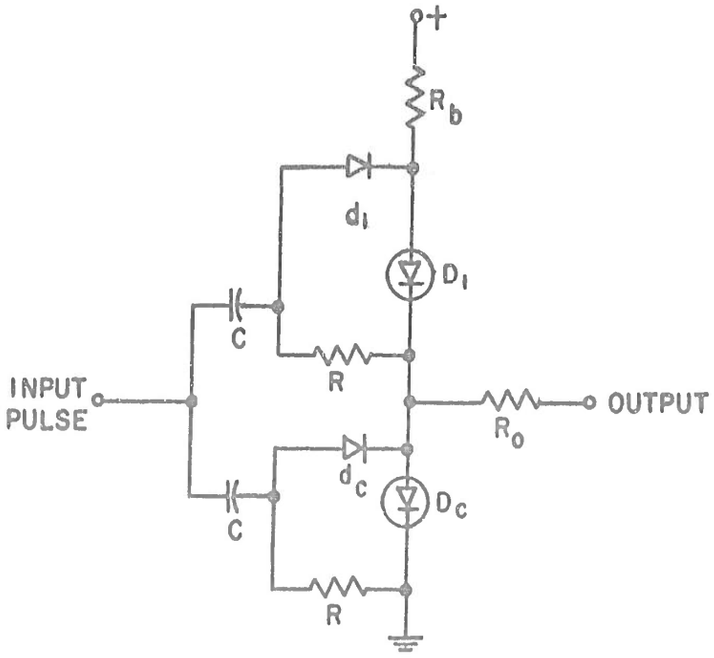
In general the clocked stages are much more satisfactory for use as shift registors but the difficulty of generating the clock signals limits their usefulness at frequencies of 1 mc or above.

The asynchronous flip-flops are primarily useful as scalers or counters. Their speed of operation is limited by the propagation time of the circuitry. Test circuits which were built had propagation times of about 40 nonoseconds per stage which limited the maximum frequency to 25 mc or less.

The low power consumption and small size of tunnel diodes offer advantages in these applications. For these reasons development effort should be continued on this circuitry.



(a)



(b)

FIG. 13 Asynchronous Tunnel Diode Flip-Flops

VIII. Pulse Generators

The development of tunnel diode and other high speed memories has caused a need for fast high-power pulses. The widths required are less than can readily be generated by transistor circuits. High current tunnel diodes offered the possibility of obtaining the high speed with the required output power.

The monostable configuration was chosen for the circuit as shown in Fig. 12b. The reasons for this were (1) the pulse width was determined by the circuit elements; (2) the turn on gain was high; and (3) no turnoff signal was required. In addition, the timing inductor afforded an excellent means of coupling signal out through a secondary winding.

A 100 ma gallium arsenide tunnel diode was obtained from ASL for test purposes. Rough calculations indicated that an inductance of about 50 nanohenries was required for a 20 nanosecond pulse. This inductance was to be supplied by means of a transformer.

Current pulses of approximately 80 milliamperes could be obtained but the minimum width was 40 microseconds. Experimental work indicated that the major problem was in the transformer which could not be adequately characterized at the frequencies involved. Several different designs were tried, some according to the methods of Ruthroff[9]. Transmission lines were also used for pulse timing and coupling out; however, the reflections encountered were so severe as to make this approach impractical.

A more detailed analysis of the operation of the monostable circuit is available but its usefulness is questionable due to the difficulty of adequately characterizing the transformer. As a result this method of using high current diodes was abandoned.

The second approach to building high power drivers is to use a low current (10 ma) tunnel diode to generate the basic pulse and use a transistor amplifier to raise the power level. The frequency response of the transistor circuitry is improved by biasing for linear operation.

No difficulty is encountered in generating 5 nanosecond pulses with a low current diode in a monostable circuit. The greatest problem has been to obtain adequate power gain in the transistor amplifier. The highest gain obtained to date has been 17 db with an amplifier output of 80 milliwatts.

The low gain has been due to the frequency response of the amplifier. Too much of the pulse energy is above the cutoff frequency of the amplifier. Another approach has now been taken in which a high current tunnel diode is used to drive a pair of common base stages. Power gain is obtained in the amplifier and the desired output level is obtained by means of a transformer. The transformer is more satisfactory in this application because it operates at a higher impedance level.

The high current tunnel diode is used in order to reduce the number of stages required in the amplifier. It is not operated in the monostable mode but rather is used to sharpen the rise and fall of the input signal which may be a sine wave. The circuit is shown in Fig. 14. When used in this mode the input signal determines the width of the pulse.

The frequency response of the transistor amplifier is greatly improved by using the transistors in the common base mode. It is not possible to obtain current amplification in the transistor but power gain is available. The use of transformers for interstage coupling and on the output allows current amplification and the proper output to be obtained. The transformers are designed according to Ruthroff's article. Output signals of 6 volts across 51 ohms have been obtained with pulse widths of 10 nonoseconds.

Either of these two methods appears capable of generating the fast high-power pulses. The use of the low current diode requires more amplification but the monostable circuit is often more convenient. The choice of circuits will depend upon the application.

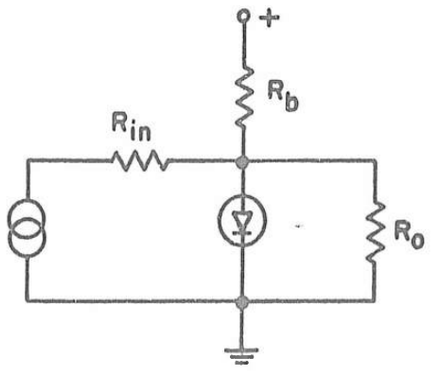


FIG. 14 Current Tunnel Diode Pulse Generator

IX. Tunnel Diode Memory

The high switching speeds and relatively low power requirements of tunnel diodes make them attractive for use in high speed memories. Many memory configurations have been suggested.

There are two main types of tunnel diode memory elements, those that use the bistable switching characteristic of the diode directly and those that use it to cause an oscillating or non-oscillating condition of a second diode. There are further subdivisions such as destructive or non-destructive readout coincident current, linear select, etc. The approach taken here has been to use the switching characteristic because the writing process and biasing appear to be less critical.

The non-destructive mode of operation has been selected because it is both simpler and faster. Writing is required only when the stored information is changed. In addition the non-destructive read operation involves pulsing the diode so that it momentarily moves a small distance along its nonlinear characteristic. This is a faster operation than switching the tunnel diode between two stable points on its nonlinear characteristic. Linear selection is used to eliminate the half-select noise of coincident current memories and thus reduce the noise problem.

The particular memory configuration chosen is shown in Fig. 15. The basic memory element is the "Goto pair". The two primary advantages of this memory are that the tolerances on the diodes used as memory

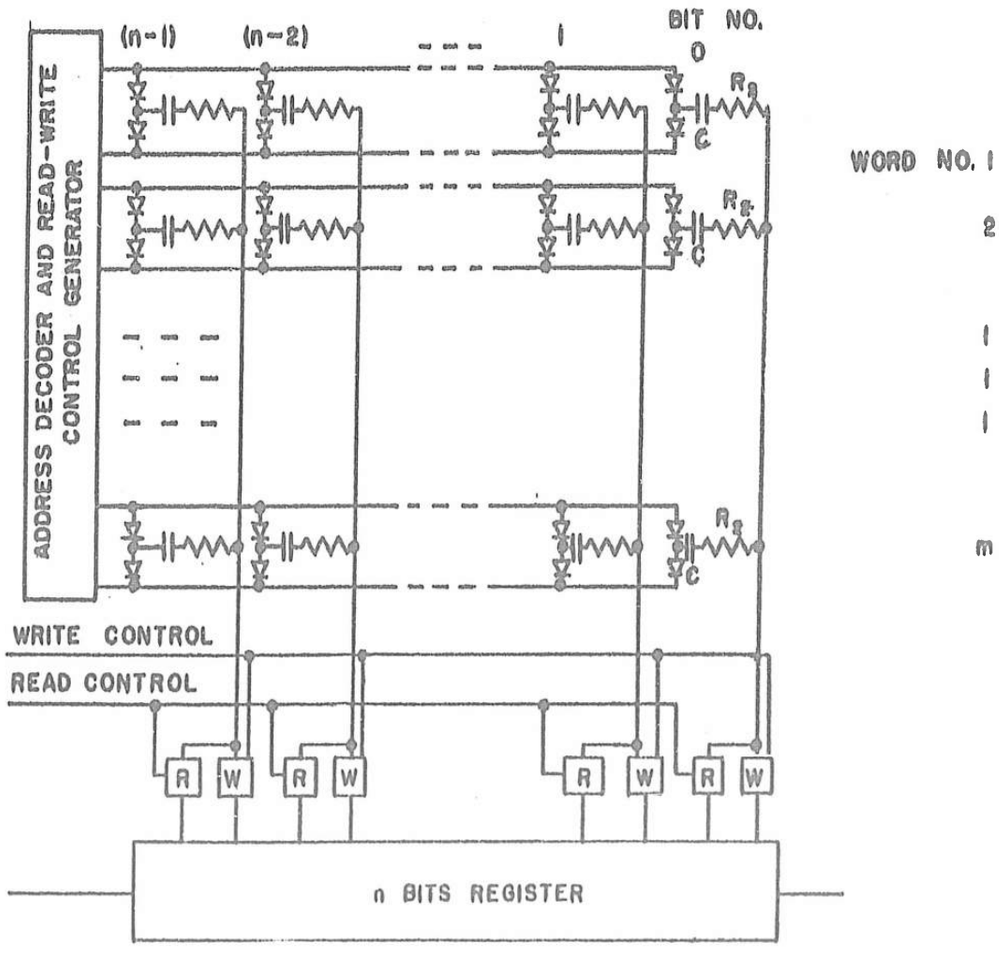


Figure 15. Tunnel Diode Memory Array Using "Goto-Pair" Elements

elements are not critical and the "one" and "zero" outputs are of opposite polarity, improving the "one" - "zero" ratio.

The memory elements of a word are supplied by a pair of word control lines. A d-c supply, complementary with respect to ground, is provided to hold the diodes in a given state. All control signals occur on the word lines. Corresponding bits are connected to common digit lines which carry the input and output information.

The principles of operation are illustrated in Fig. 16. The d-c bias is such that only one diode can be in region 3, the high voltage state, while the other must be in region 1, the low voltage state. To insert information, the write control generator reduces the bias voltage across the diodes of a word such that both diodes of a storage element are reset to region 1. Prior to the recovery of the d-c bias, a digit line write signal in the form of a positive or negative current pulse is applied to the common point through capacitor C. The write signal acts as a prejudice current and causes one of the diodes to switch into region three just as in the conventional majority logic circuit.

To interrogate the storage elements of a word, the read control generator reduces the d-c bias by a small amount. This reduction in bias is not sufficient to reset the diodes but it does produce an output signal the polarity of which is determined by the stored information. This reduction in voltage is in the form of a pulse and it appears across each pair of diodes in a word. The voltage

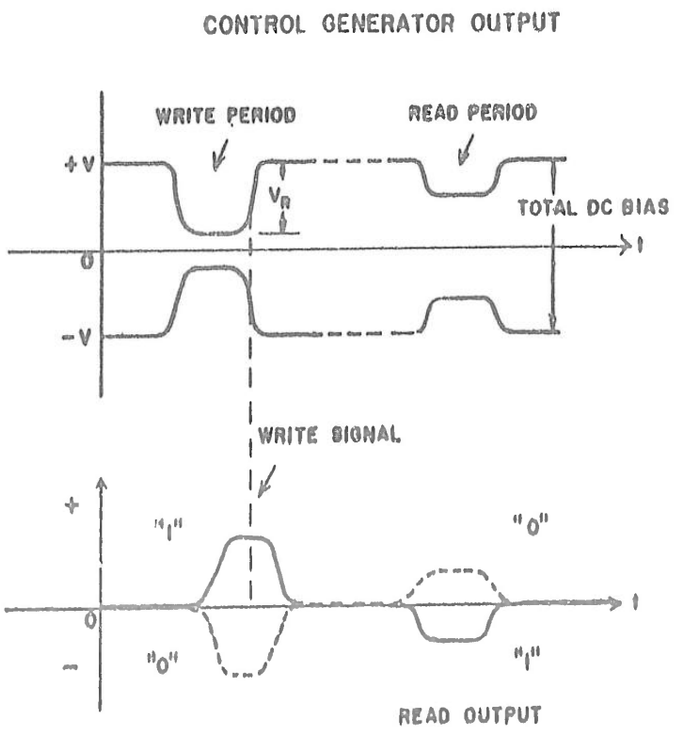


Figure 16. Operation Principle of the Memory Array Using "Goto-Pair" Elements

across the diode in the low voltage state can change only slightly so most of the reduction in bias must appear across the diode in the high voltage state. As a result, the read control signal is transferred to the output point across the diode in the low voltage state with little attenuation. Because of this the polarity of the output signal is opposite that of the write signal.

For "Goto pair" circuits with low fan in and fan out numbers the tolerance requirements are not severe. In the memory the maximum fan in and fan out is one. Consequently tolerances are not a problem. The writing signal must be sufficient to overcome any difference in the diode peak currents under the worst conditions. In practice circuit and diode tolerances of $\pm$ 10% will not adversely affect operation.

The drivers provide the signals required to operate the memory. The read and write control signals occur on the control lines. These control lines must have low impedance os that they are effectively voltage sources. This is necessary so that the tolerance problem does not arise and also to prevent oscillations. The required driver current is a function of the number of bits in a word since the memory elements are in parallel. The read control driver need not provide as much current as the write control driver because of its smaller signals. In addition because of the desire for high-speed operation these pulses must be very narrow.

The amplitudes of the pulses are determined by the diode characteristics. For germanium tunnel diodes the d-c bias voltages are between ± 200 mv and ± 225 mv for proper operation. This calls for a write control signal of approximately 200 mv and a read control signal of approximately ± 50 mv so that these signals are at low voltage levels as well as low impedance levels.

The first attempt to generate these signals was by means of transistor blocking oscillators with transformers used to generate the low voltages and low impedances. The circuit is shown in Fig. 17. This circuit was totally unsatisfactory for a number of reasons. First, the method of coupling the signals into the control lines was poor and wasted too much power. Second, the pulses generated by the transistor circuits were too wide. The third reason was that satisfactory pulse transformers could not be obtained. The low voltages could be generated but the impedance level was far too high due to poor coupling in the rransformers. This approach was dropped and a study was made of tunnel diode pulse generators as reported on in section VIII.

The proposed use of the new pulse generators is to generate the pulses at reasonably high voltages with resistive voltage dividers reducing the signals to the desired level. This method is wasteful of power but it holds more promise than the use of transformers.

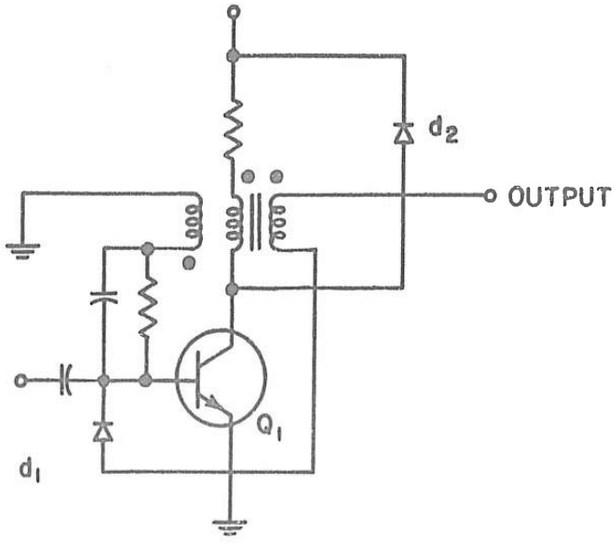


FIG. 17 Transistor Blocking Oscillator

The situation is somewhat different for the write driver. This driver provides the write or prejudice current along the digit lines. The write signal is applied to the memory elements through a capacitor and a resistor which can be made as large as necessary. As a result the write signal can be a high voltage signal. In addition, the load on the write driver is constant and resistive, consequently it is not necessary to dump current into a sink to provide the proper output characteristics. The primary problem with the write generator is to provide signals that are sufficiently narrow. In addition, the write generator must be able to provide signals of either polarity.

The original method of building the write generator was to use a pair of transistor blocking oscillators. The one that was triggered was determined by the information to be stored in the memory. The pulses from the blocking oscillator were too wide. The need for narrower pulses led to the use of tunnel diodes as the basic pulse generators followed by linear amplifiers as described in section VIII. Further improvements can be made by using only one amplifier circuit and combining the outputs of the two pulse generators at the amplifier input. This has not been done yet and it does not appear to be straight-forward but it should be possible.

The pulse generators are shown in Fig. 18. The circuit is in the form of a tunnel diode pulse gate. One of the inputs is the d-c gating signal from the input register and the other is the clock pulse. Only the diode with both inputs present will be triggered.

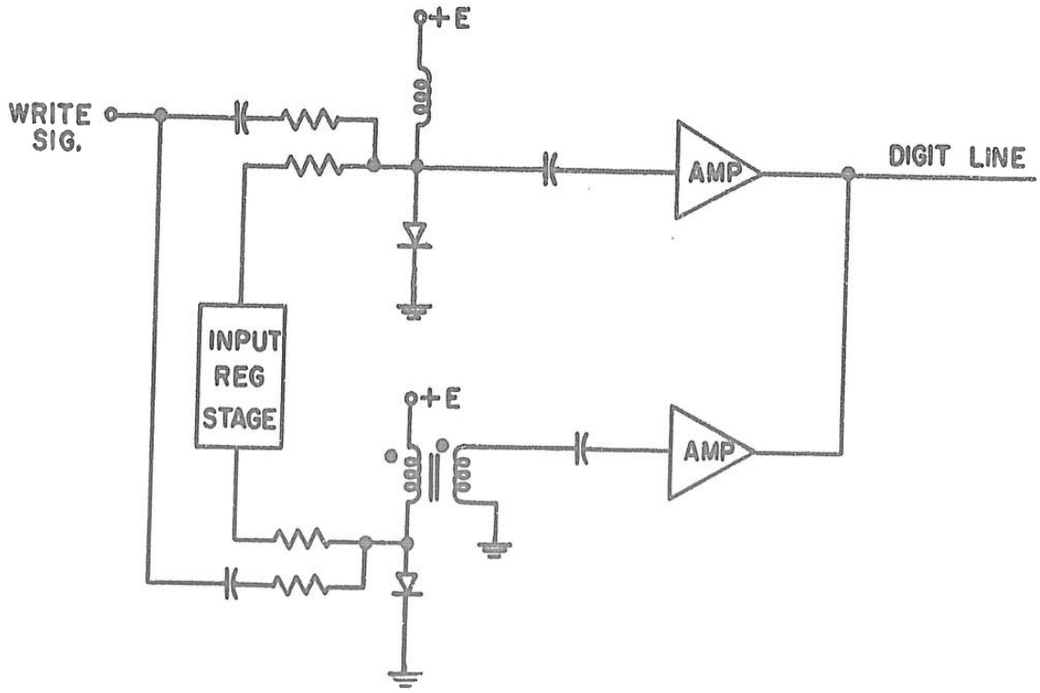
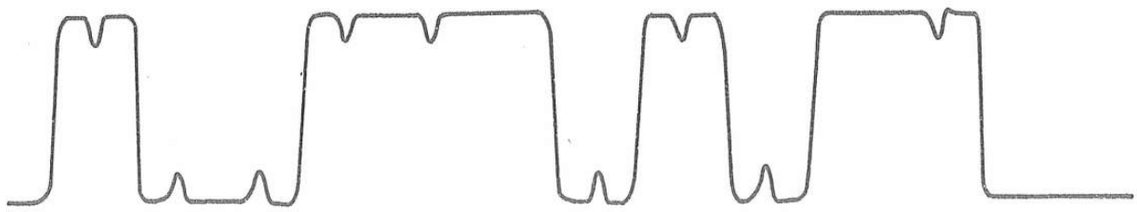


FIG. 18 Write Pulse Generators

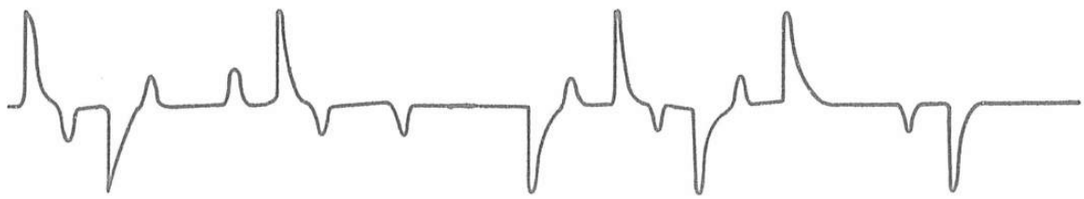
This circuit has been operated successfully.

An additional problem has been encountered in operating the memory. The original control lines were open wires placed in close proximity to the tunnel diodes. At the frequencies of operation the distributed parameters of this line caused it to behave as a transmission line of unknown characteristics. Experimentally determined terminations were found to improve operation. It has become apparent that much more work will have to be done on this problem.

The output of a memory element appears at the common point of the diode pair and is caused by the read control signal as shown in Fig. 19. The effect of the write signals also appears at this point. The amplitude of the read out signal is approximately 50 mv while the d-c level swings $\pm$ 200 mv due to writing. As can be seen from Fig. 15 the memory output is capacitively coupled (for isolation purposes) and the d-c level cannot be sensed. The sensed information is in the form of pulses with the d-c level eliminated. The output signal must be differentiated strongly to prevent the writing transients from interfering with the read signal. The output signal is attenuated due to the loading of all the other memory elements on the same digit line. This attenuation limits the number of words that may be connected in paralled. The attenuation can be calculated with the help of Fig. 20.



(a) OUTPUT AT COMMON POINT



(b) DIFFERENTIATED OUTPUT

FIG. 19

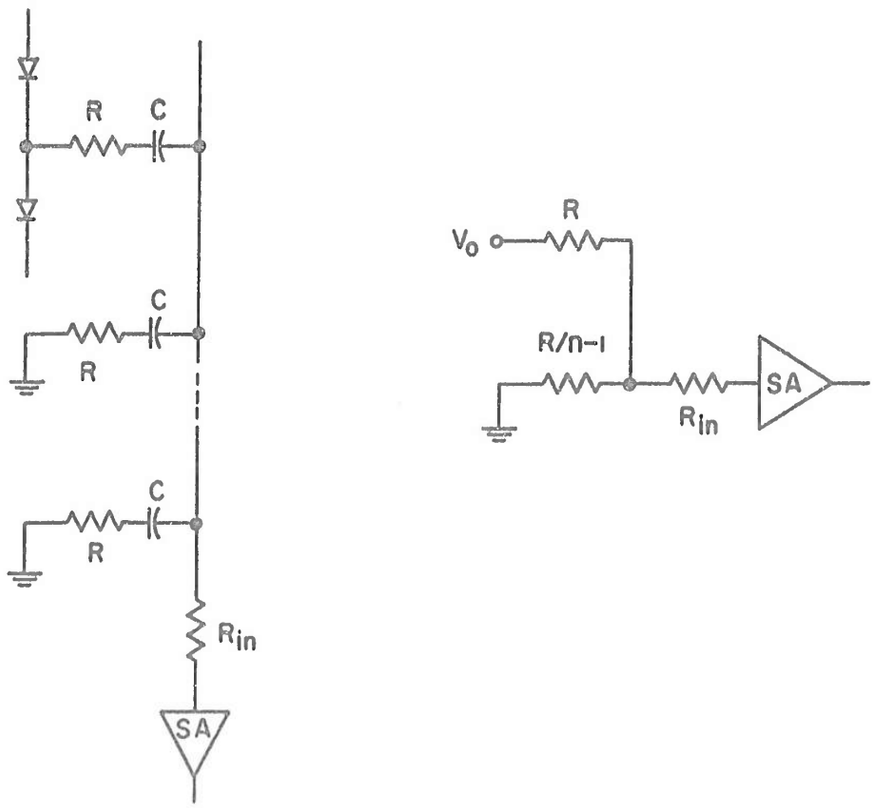


FIG. 20 Circuit for Loading Calculation

The signal voltage at the input to the sense amplifier is given by

$$V_i = \frac{V_o(X)}{1 + nx}$$

where

$X = R_i/R$

R_i = Input impedance of sense amplifier

V_o = Output at common point of diodes

n = Number of bits in parallel on digit line

If $nx >> 1$ as it would be in practice then $V_i \approx V_o/n$. For an output signal of 50 mv, the voltage at the input to the amplifier is only 5 mv for 10 words in parallel. Consequently, the sense amplifier must provide considerable amplification if its output is to be capable of setting the output register.

In addition the sense amplifier also contains the strobe circuitry required to separate the output signal from the writing transients and other noise. The original strobe circuit consisted of a four diode gate as shown in Fig. 21. The operation of the gate was such that for an input signal to noise ratio of 1/4, the signal to noise ratio of the output was 4. This improvement is made possible by the time separation of signal and noise. The strobe gate was operated by a blocking oscillator triggered from the read control driver.

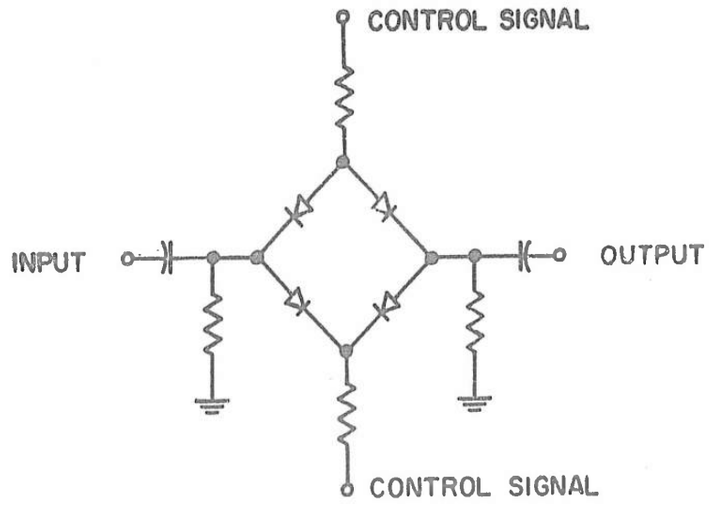


FIG. 21 Diode Strobe Gate

The original sense amplifier consisted of 2N706 transistors used in a grounded emitter video amplifier configuration. It is anticipated that a different configuration may be used to obtain greater amplification. In addition the possibility of using a tunnel diode for both strobing and gain as shown in Fig. 22 will be investigated.

The original memory was constructed with open wire control lines and transistor drivers. The memory consisted of ten words of ten bits each. The entire memory was never operated at once; however, a twenty bit word was made using two ten bit words. This memory was operated non-destructively with a read cycle of 100 nanoseconds and a write cycle of 100 nanoseconds. This allowed reading or writing alone at a 10 mc rate. Operation was critical at these speeds because of drivers and control line characteristics. The memory work was suspended while the driver development program was undertaken.

With satisfactory drivers now available the memory work is being resumed. The major problem is now one of providing proper control line characteristics. Two methods have been suggested. One involves the use of strip lines and the other involves the use of tapped coaxial cable. In any case it appears that open wire lines can no longer be used.

The objective of the future work is to demonstrate the practicality of a memory of ten words or more, of fifty bits or more each. The memory is to operate at as high a rate as possible with the original goal set at 25 mc.

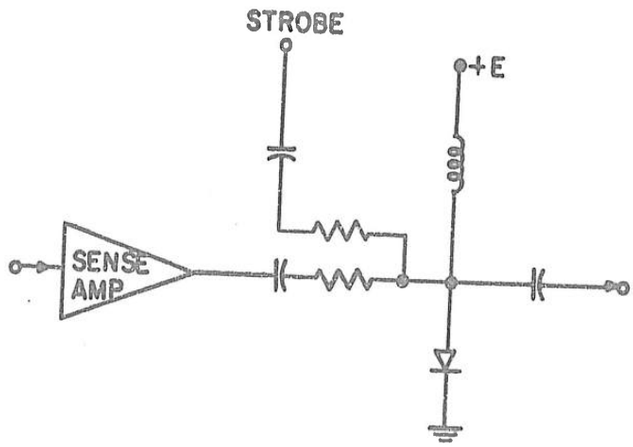


FIG. 22 Strobe Gate

X. Active R-C Filters

A study was made of the possibility of designing band pass filters using R-C elements and tunnel diodes only. Analysis showed that for stable operation the limitations on the critical frequencies of networks using R-C elements and tunnel diodes are the same as those for networks using R-C elements alone. The poles of the transfer junction are restricted to the real frequency axis. With these restrictions a large number of elements is required to achieve a reasonably narrow band amplifier. About 400 elements would be required to achieve a Q of 5.

There is one important difference between circuits with or without tunnel diodes. It is possible to obtain large power gains when tunnel diodes are used.

A simple bandpass filter was designed and built using R-C elements and tunnel diodes. The measured power gain was 16.4db and the Q was 0.5.

XI. Continuing Investigations

Work is to be continued in vapor (halogen) transport to grow intermetallic compounds and mixed compounds. Attempts are to be made to grow epitaxially GaAs tunnel junctions and to study the reliability problem on grown junctions as well as alloyed junctions. The possibility of epitaxially growing junctions in mixed intermetallic compounds is to be attempted. Here there is the possibility of varying the bandgap on the two sides of the junction (tunnel junction or normal junction). Potentially interesting mixed crystals for tunnel diodes are to be considered.

In addition, the various tunnel diode computer and logic studies and problems described in sections VI through X are to be continued.

Scientists and technicians who contributed to the work reported:

Staff	Technicians
N. Holonyak, Jr.	S. Bevacqua
E. Fisch	M. McCarthy

References

1. T. A. Longo, "On the Nature of the Maximum and Minimum Currents in Germanium Tunnel Diodes," Bull. Am. Phys. Soc., Series II, vol. 5, p. 160, March, 1960.
2. N. Holonyak, Jr., D. C. Jillson, and S. F. Bevacqua, "Silicon, Arsenic, Whiskers, and Tunnel Diodes," A.I.M.E. Boston Meeting Aug. 29, 1960, (to be published by Interscience Pub., N.Y.).
3. L. Esaki, "New Phenomenon in Narrow Ge p-n Junctions," Phys. Rev., vol. 109, p. 603, Jan., 1958.
4. N. Holonyak, Jr. and I. A. Lesk, "Gallium Arsenide Tunnel Diodes," Proc. I.R.E., vol. 48, p.1405, August, 1960.
5. J. C. Marinace, "Epitaxial Vapor Growth of Ge Single Crystals in a Closed-Cycle Process," I.B.M. Jour., vol. 4, p.248, July, 1960.
6. U. S. Davidsohn, Y. C. Hwang, and G. B. Ober, "Designing With Tunnel Diodes," Electronic Design, Feb. 3, and Feb. 17, 1960.
7. U. S. Davidsohn, "Notes on the Measurement of Tunnel Diode Junction Capacitance."
8. Chow, W. F., "Tunnel Diode Digital Circuitry", 1960 International Solid State Circuits Conference.
9. Ruthroff, C. L., "Some Broad Band Transformer", pp. 1337-1342, Proc. I.R.E., Aug. 1959.

FIGURES